10.1 Comparing Two Proportions

(Day 1)

vocab

examples

Objectives

B-Baller

Is it harder to shoot free-throws with distractions? To investigate, a basketball player went to the gym and shot 20 free-throws. Ten of the free-throws were shot without any distractions and the other 10 were shot with his friends trying everything they could to distract him. The order of the 20 shots was determined at random.

Why was it important that the order of the shots was determined at random, rather than doing all of one type of shot before the other type of shot?

B-Baller

The player made 5/10 (50%) of his shots in the distraction-free environment and only 3/10 (30%) of his shots in the environment with distractions, for a difference of 50% – 30% = 20%.

Identify two plausible explanations for why the shooter performed better in the distraction-free environment.

B-Baller Simulation

20 notecards:
- 8 - YES
- 12 - NO

Shuffle and randomly pick 10

Use your stickers to dot plot your results

sampling distribution of $\hat{p}_1 - \hat{p}_2$

Describes the possible values of $\hat{p}_1 - \hat{p}_2$ and how often they will occur

Choose an SRS of size n_1 from Population 1 with proportion of successes p_1 and an independent SRS of size n_2 from Population 2 with proportion of successes p_2.

Shape: When n_1p_1, $n_1(1 - p_1)$, n_2p_2, and $n_2(1 - p_2)$ are all at least 10, the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is approximately Normal.

sampling distribution of $\hat{p}_1 - \hat{p}_2$

Center: The mean of the sampling distribution is $p_1 - p_2$. That is, the difference in sample proportions is an unbiased estimator of the difference in population proportions.

Spread: The standard deviation of the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

as long as the 10% condition is met.

* on formula sheet (back)

The snack that smiles back :-)

Your teacher brings two bags of colored goldfish crackers to class. She tells you that Bag 1 has 25% red crackers and Bag 2 has 35% red crackers. Each bag contains more than 500 crackers. Using a paper cup, your teacher takes an SRS of 50 crackers from Bag 1 and a separate SRS of 40 crackers from Bag 2. Let $\hat{p}_1 - \hat{p}_2$ be the difference in the sample proportions of red crackers.

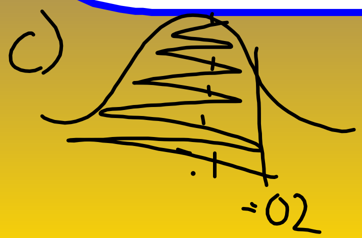
(a) What is the shape of the sampling distribution of $\hat{p}_1 - \hat{p}_2$? Why?

(b) Find the mean and standard deviation of the sampling distribution. Show your work.

(c) Find the probability that $\hat{p}_1 - \hat{p}_2$ is less than or equal to −0.02. Show your work.

(d) Based on your answer to c, would you be surprised if the difference in the proportion of red crackers in the two samples was $\hat{p}_1 - \hat{p}_2 = -0.02$?

c)



normalcdf (-1E99, -.02, -.1, .097) = .795

low up mean sd

10.1 Comparing Two Proportions

(Day 2)

vocab

examples

Two Sample z Interval for a Difference between Two Proportions

Check three conditions:

Random: The data are produced by a random sample of size n_1 from Population 1 and a random sample of size n_2 from Population 2 or by two groups of size n_1 and n_2 in a randomized experiment

Normal: $np \geq 10$ and $n(1-p) \geq 10$ for both $n_1\hat{p}_1$ and $n_2\hat{p}_2$

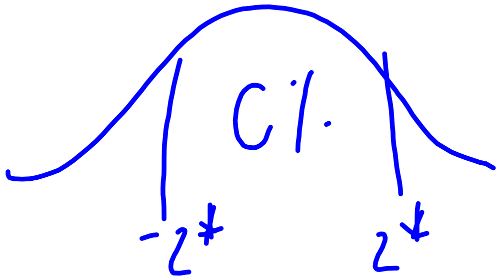
Independent: Both the samples or groups themselves and the individual observations in each sample or group are independent. The 10% condition must be met.

Two Sample z Interval for a Difference between Two Proportions

When the Random, Normal, and Independent conditions are met, an approximate level C confidence interval for $\hat{p}_1 - \hat{p}_2$ is:

$$(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

where z* is the critical value for the standard Normal curve with area C between -z* and z*



Social Networking

As part of the Pew Internet and American Life Project, researchers conducted two surveys in late 2009. The first survey asked a random sample of 800 U.S. teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 U.S. adults. In these two studies, 73% of teens and 47% of adults said that they use social-networking sites. Use these results to construct and interpret a 95% confidence interval for the difference between the proportion of all U.S. teens and adults who use social-networking sites.

Social Networking

As part of the Pew Internet and American Life Project, researchers conducted two surveys in late 2009. The first survey asked a random sample of 800 U.S. teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 U.S. adults. In these two studies, 73% of teens and 47% of adults said that they use social-networking sites. Use these results to construct and interpret a 95% confidence interval for the difference between the proportion of all U.S. teens and adults who use social-networking sites.

STATE: We want to estimate $p_1 - p_2$ at a 95% confidence level.
p_1 is the proportion of teens who use social media. p_2 is the proportion of adults who use social media.

Social Networking

As part of the Pew Internet and American Life Project, researchers conducted two surveys in late 2009. The first survey asked a random sample of 800 U.S. teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 U.S. adults. In these two studies, 73% of teens and 47% of adults said that they use social-networking sites. Use these results to construct and interpret a 95% confidence interval for the difference between the proportion of all U.S. teens and adults who use social-networking sites.

PLAN: Random: yes, both samples are random

Independent: assuming at least 8000 teens & 22530 adults in US

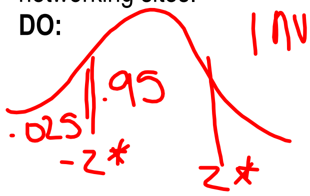
Normal: $800(.73) \geq 10$ ✓
$800(1-.73) \geq 10$ ✓ yes!
$2253(.47) \geq 10$ ✓
$2253(1-.47) \geq 10$ ✓

We use a two sample z interval for $p_1 - p_2$

Social Networking

As part of the Pew Internet and American Life Project, researchers conducted two surveys in late 2009. The first survey asked a random sample of 800 U.S. teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 U.S. adults. In these two studies, 73% of teens and 47% of adults said that they use social-networking sites. Use these results to construct and interpret a 95% confidence interval for the difference between the proportion of all U.S. teens and adults who use social-networking sites.

DO:



invNorm(.025, 0, 1) → $\pm 1.96 = \pm z^*$

(area below, μ, σ)

$$(.73-.47) \pm 1.96\sqrt{\frac{.73(1-.73)}{800}+\frac{.47(1-.47)}{2253}}$$

$$.26 \pm 1.96\sqrt{.0003569}$$

$$.26 \pm .037 \quad (.223, .297)$$

Social Networking

As part of the Pew Internet and American Life Project, researchers conducted two surveys in late 2009. The first survey asked a random sample of 800 U.S. teens about their use of social media and the Internet. A second survey posed similar questions to a random sample of 2253 U.S. adults. In these two studies, 73% of teens and 47% of adults said that they use social-networking sites. Use these results to construct and interpret a 95% confidence interval for the difference between the proportion of all U.S. teens and adults who use social-networking sites.

CONCLUDE:

We are 95% confident that the interval from .223 to .297 contains true difference of proportion of teens & adults who use social media. This suggests that teens use social media more.

On the calculator...

Gun Laws

Have opinions changed about gun control? Gallup regularly asks random samples of U.S. adults their opinion on a variety of issues. In a poll of 1011 U.S. adults in January 2013, 38% responded that they “were dissatisfied with the nation’s gun laws and policies, and want them to be stricter.” In a similar poll of 1011 adults in January 2012, only 25% agreed with this statement. Use the results of these polls to construct and interpret a 90% confidence interval for the change in the proportion of U.S. adults who would agree with the statement about gun laws.

Gun Laws

Have opinions changed about gun control? Gallup regularly asks random samples of U.S. adults their opinion on a variety of issues. In a poll of 1011 U.S. adults in January 2013, 38% responded that they "were dissatisfied with the nation's gun laws and policies, and want them to be stricter." In a similar poll of 1011 adults in January 2012, only 25% agreed with this statement. Use the results of these polls to construct and interpret a 90% confidence interval for the change in the proportion of U.S. adults who would agree with the statement about gun laws.

STATE: We want to estimate the difference of p_1-p_2 at a 90% confidence level. p_1 is the proportion of US adults who said they were dissatisfied with gun laws in 2013 and p_2 is the proportion of adults who had the same opinion in 2012.

Gun Laws

Have opinions changed about gun control? Gallup regularly asks random samples of U.S. adults their opinion on a variety of issues. In a poll of 1011 U.S. adults in January 2013, 38% responded that they "were dissatisfied with the nation's gun laws and policies, and want them to be stricter." In a similar poll of 1011 adults in January 2012, only 25% agreed with this statement. Use the results of these polls to construct and interpret a 90% confidence interval for the change in the proportion of U.S. adults who would agree with the statement about gun laws.

PLAN: Normal: $1011(.38) \geq 10$ and $1011(1-.38) \geq 10$

$1011(.25) \geq 10$ and $1011(1-.25) \geq 10$ yes!

Independent : assuming at least 10110 adults in the US

Random: yes, both samples were randomly collected

We will use a two sample z interval for p_1-p_2

Gun Laws

Have opinions changed about gun control? Gallup regularly asks random samples of U.S. adults their opinion on a variety of issues. In a poll of 1011 U.S. adults in January 2013, 38% responded that they "were dissatisfied with the nation's gun laws and policies, and want them to be stricter." In a similar poll of 1011 adults in January 2012, only 25% agreed with this statement. Use the results of these polls to construct and interpret a 90% confidence interval for the change in the proportion of U.S. adults who would agree with the statement about gun laws.

DO: on the calculator... 2-PropZInt x1: 384 n1: 1011 x2: 253 n2: 1011

c-level: .9

(.09592, .16323)

Gun Laws

Have opinions changed about gun control? Gallup regularly asks random samples of U.S. adults their opinion on a variety of issues. In a poll of 1011 U.S. adults in January 2013, 38% responded that they "were dissatisfied with the nation's gun laws and policies, and want them to be stricter." In a similar poll of 1011 adults in January 2012, only 25% agreed with this statement. Use the results of these polls to construct and interpret a 90% confidence interval for the change in the proportion of U.S. adults who would agree with the statement about gun laws.

CONCLUDE: We are 90% confident that the interval from .09592 to .16323 contains the true difference in proportions of US adult's opinions from 2013 to 2012. So, people are more concerned about gun safety now than they were a year ago.

10.1 Comparing Two Proportions

(Day 3)

vocab

examples

pooled sample proportion

combine both samples together

$$\hat{p}_C = \frac{\text{count of successes in both samples combine}}{\text{count of individuals in both samples combine}} = \frac{X_1 + X_2}{n_1 + n_2}$$

Test Statistic

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\textit{standard deviation of statistic}}$$

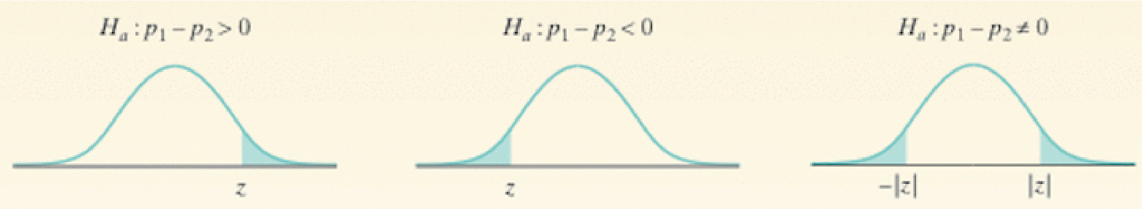
Two Sample z Test for the Difference between two proportions

Suppose the Random, Normal, and Independent conditions are met. To test the hypothesis $H_0: p_1 - p_2 = 0$, first find the pooled proportion $\hat{p}_C$ of successes in both samples combined. Then compute the z statistic:

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1-\hat{p}_C)}{n_1} + \frac{\hat{p}_C(1-\hat{p}_C)}{n_2}}}$$

Two Sample z Test for the Difference between two proportions

Find the P-value by calculating the probability of getting a z statistic this large or larger in the direction specified by the alternative hypothesis H_a:



Two Sample z Test for the Difference between two proportions

Check three conditions:

Random: samples are random from both

Normal: $np \geq 10$ and $n(1-p) \geq 10$ for both n_1p_1 and n_2p_2

Independent: Both the samples or groups themselves and the individual observations in each sample or group are independent. The 10% condition must be met.

Use 4 step process:

STATE, PLAN, DO, CONCLUDE

On the calculator...

It is encouraged to use the calculator for the tests in this chapter!

Use 2-PropZTest

Hungry Children

Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 students included 26 who had not had breakfast. More than 1500 students attend each school. Do these data give convincing evidence of a difference in the population proportions? Carry out a significance test at the $\alpha = 0.05$ level to support your answer.

Hungry Children

Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 students included 26 who had not had breakfast. More than 1500 students attend each school. Do these data give convincing evidence of a difference in the population proportions? Carry out a significance test at the α = 0.05 level to support your answer.

STATE:

$H_0: p_1 - p_2 = 0$

$H_a: p_1 - p_2 \neq 0$

$\alpha = .05$

p_1 = proportion of students who didn't eat b-fast @ school 1.

p_2 = proportion of students who didn't eat b-fast @ school 2.

Hungry Children

Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 students included 26 who had not had breakfast. More than 1500 students attend each school. Do these data give convincing evidence of a difference in the population proportions? Carry out a significance test at the α = 0.05 level to support your answer.

PLAN:

Random - yes, both random

Independent - assuming at least 800 students @ school 1 & 1500 @ school 2

Normal - $80(.24) \geq 10$ $\quad 150(.17) \geq 10$

$80(1-.24) \geq 10$ $\quad 150(1-.17) \geq 10$

We will use a 2 sample z test for $p_1 - p_2$!!!!

Hungry Children

Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 students included 26 who had not had breakfast. More than 1500 students attend each school. Do these data give convincing evidence of a difference in the population proportions? Carry out a significance test at the α = 0.05 level to support your answer.

DO:

2-PropZTest

x_1: 19 n_1: 80

x_2: 26 n_2: 150

$p_1 \neq p_2$

z = 1.17

P-VALUE = .24

Hungry Children

Researchers designed a survey to compare the proportions of children who come to school without eating breakfast in two low-income elementary schools. An SRS of 80 students from School 1 found that 19 had not eaten breakfast. At School 2, an SRS of 150 students included 26 who had not had breakfast. More than 1500 students attend each school. Do these data give convincing evidence of a difference in the population proportions? Carry out a significance test at the α = 0.05 level to support your answer.

CONCLUDE:

.24 ≥ .05 → we fail to Reject Ho → we cannot conclude that the difference in proportions of students who didn't eat b-fast @ school 1 & school 2 does not equal 0. In other words, we can't conclude that the prop. are different btwn school 1 & school 2. The results are NOT statistically sig. @ 5% level

Teens Going Deaf?

In a study of 3000 randomly selected teenagers in 1988–1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005–2006, 19.5% showed some hearing loss.(These data are reported in Arizona Daily Star, August 18, 2010)

Do these data give convincing evidence that the proportion of all teens with hearing loss has increased?

Teens Going Deaf?

In a study of 3000 randomly selected teenagers in 1988–1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005–2006, 19.5% showed some hearing loss.(These data are reported in Arizona Daily Star, August 18, 2010)

Do these data give convincing evidence that the proportion of all teens with hearing loss has increased?

STATE: H_o: $p_1 - p_2 = 0$ H_a: $p_1 - p_2 < 0$ $\alpha = .05$

p_1 is the proportion of teens in 1988-2004 who showed some hearing loss

p_2 is the proportion of teens in 2005-2006 who showed some hearing loss

Teens Going Deaf?

In a study of 3000 randomly selected teenagers in 1988–1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005–2006, 19.5% showed some hearing loss.(These data are reported in Arizona Daily Star, August 18, 2010)

Do these data give convincing evidence that the proportion of all teens with hearing loss has increased?

PLAN: Random - yes, both random

Independent - assuming at least 30,000 teens from 1988-1994 and at least 1800 teens from 2005 -2006

Normal - $3000(.15)\geq 10$ $1800(.195)\geq 10$

$3000(1-.15)\geq 10$ $3000(1-.195)\geq 10$

We will use a two sample z test for p_1-p_2

Teens Going Deaf?

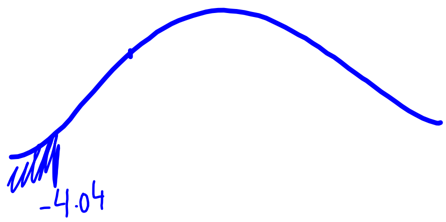
In a study of 3000 randomly selected teenagers in 1988–1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005–2006, 19.5% showed some hearing loss.(These data are reported in Arizona Daily Star, August 18, 2010)

Do these data give convincing evidence that the proportion of all teens with hearing loss has increased?

DO: 2-propZTest

x1: 450 n1: 3000 x2: 351 n2: 1800

$p_1<p_2$ z=-4.04 p-VALUE = .000026



Teens Going Deaf?

In a study of 3000 randomly selected teenagers in 1988–1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005–2006, 19.5% showed some hearing loss.(These data are reported in Arizona Daily Star, August 18, 2010)

Do these data give convincing evidence that the proportion of all teens with hearing loss has increased?

CONCLUDE: .000026 < .05 --> We reject H_0 --> We can conclude that the difference in proportions of teens who showed some hearing loss from 1988-1994 to 2005-2006 is less than 0. In other words, there is convincing evidence to show that the proportion of all teens with hearing loss has increased. The results are statistically significant at the 5% level.

10.1 Comparing Two Proportions

(Day 4)

vocab

examples

Notation

H_0: $p_1 - p_2 = 0$ means the same as.... H_0: $p_1 = p_2$

H_a: $p_1 - p_2 \neq 0$ means the same as... H_a: $p_1 \neq p_2$

H_a: $p_1 - p_2 > 0$ means the same as... H_a: $p_1 > p_2$

H_a: $p_1 - p_2 < 0$ means the same as... H_a: $p_1 < p_2$

Random

The random condition is satisfied for a randomized experiment when subjects are randomly assigned their treatments

Cash for Quitters

In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking?

Cash for Quitters

In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking?

STATE: H_0: $p_1 - p_2 = 0$ H_a: $p_1 - p_2 > 0$ $\alpha = .05$

p1 is the proportion of smokers who quit with a financial reward. p2 is the proportion of smokers who quit with traditional methods.

Cash for Quitters

In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking?

PLAN: Random - yes, random assignment

Independent - assuming at least 8780 smokers

Normal - 439 (.15) ≥ 10 439(.05)≥10

439(1-.15) ≥ 10 439 (1-.05) ≥ 10

We will use a two sample z test for p1-p2

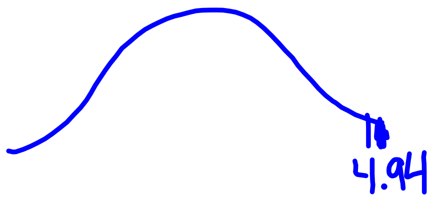
Cash for Quitters

In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking?

DO: 2-PropZTest

x1: 66 n1:439 x2: 22 n2: 439 p1>p2

z = 4.944 p-VALUE = .00000038



Cash for Quitters

In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to $750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking?

Conclude: .00000038 < .05 --> We reject the null hypothesis --> We can conclude that the difference in proportions of those who quit smoking using money as an incentive and those who quit smoking traditionally is greater than 0. Therefore, there is convincing evidence that a financial incentive helps people quit. The results are statistically significant at the 5% level.